



ENTRANCE RECEIVED APR 08 2002

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/970,638

DATE: 03/27/2002  
TIME: 13:58:40

Input Set : F:\38155-20039.txt  
Output Set: N:\CRF3\03272002\I970638.raw

3 <110> APPLICANT: Millennium Pharmaceuticals, Inc.  
 4 Meyers, Rachel  
 6 <120> TITLE OF INVENTION: 27419, A NOVEL HUMAN ARGININE-N-METHYL  
 7 TRANSFERASE AND USES THEREOF  
 10 <130> FILE REFERENCE: 38155-20039.00  
 12 <140> CURRENT APPLICATION NUMBER: US 09/970,638  
 C--> 13 <141> CURRENT FILING DATE: 2002-03-12  
 15 <150> PRIOR APPLICATION NUMBER: US 60/237,717  
 16 <151> PRIOR FILING DATE: 2000-10-05  
 18 <160> NUMBER OF SEQ ID NOS: 9  
 20 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 22 <210> SEQ ID NO: 1  
 23 <211> LENGTH: 1201  
 24 <212> TYPE: DNA  
 25 <213> ORGANISM: Homo sapiens  
 27 <220> FEATURE:  
 28 <221> NAME/KEY: CDS  
 29 <222> LOCATION: (34)...(1161)  
 31 <400> SEQUENCE: 1  
 32 attgaattta gcggccgcga attcgccctt aag atg tcg cag ccc aag aaa aga 54  
 33 Met Ser Gln Pro Lys Lys Arg  
 34 1 5  
 36 aag ctt gag tcg ggg ggc ggc gaa gga ggg gag gga act gaa gag 102  
 37 Lys Leu Glu Ser Gly Gly Gly Glu Gly Gly Thr Glu Glu  
 38 10 15 20  
 40 gaa gat ggc gcg gag cgg gag gcg gcc ctg gag cga ccc cgg agg act 150  
 41 Glu Asp Gly Ala Glu Arg Glu Ala Ala Leu Glu Arg Pro Arg Arg Thr  
 42 25 30 35  
 44 aag cgg gaa cgg gac cag ctg tac tac gag tgc tac tcg gac gtt tcg 198  
 45 Lys Arg Glu Arg Asp Gln Leu Tyr Tyr Glu Cys Tyr Ser Asp Val Ser  
 46 40 45 50 55  
 48 gtc cac gag gag atg atc gcg gac cgc gtc cgc acc gat gcc tac cgc 246  
 49 Val His Glu Glu Met Ile Ala Asp Arg Val Arg Thr Asp Ala Tyr Arg  
 50 60 65 70  
 52 ctg ggt atc ctt cgg aac tgg gca gca ctg cga ggc aag acg gta ctg 294  
 53 Leu Gly Ile Leu Arg Asn Trp Ala Ala Leu Arg Gly Lys Thr Val Leu  
 54 75 80 85  
 56 gac gtg ggc gcg ggc acc ggc att ctg agc atc ttc tgt gcc cag gcc 342  
 57 Asp Val Gly Ala Gly Thr Gly Ile Leu Ser Ile Phe Cys Ala Gln Ala  
 58 90 95 100  
 60 ggg gcc cgg cgc gtg tac gcg gta gag gcc agc gcc atc tgg caa cag 390  
 61 Gly Ala Arg Arg Val Tyr Ala Val Glu Ala Ser Ala Ile Trp Gln Gln  
 62 105 110 115

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/970,638

DATE: 03/27/2002  
TIME: 13:58:40

Input Set : F:\38155-20039.txt  
Output Set: N:\CRF3\03272002\I970638.raw

64	gcc	cg	gag	gt	gt	cg	t	tc	a	ac	gg	ct	gag	g	a	c	cg	gt	ca	c	gt	438
65	Ala	Arg	Glu	Val	Val	Arg	Phe	Asn	Gly	Leu	Glu	Asp	Arg	Val	His	Val						
66	120																125	130	135			
68	ct	cg	g	ga	cc	gt	g	act	gt	g	tt	cc	g	aa	c	ag	gt	gat	gc		486	
69	Leu	Pro	Gly	Pro	Val	Glu	Thr	Val	Glu	Leu	Pro	Glu	Gln	Val	Asp	Ala						
70																	140	145	150			
72	atc	gt	tg	ag	g	tg	tg	at	g	gc	t	tc	ct	c	ac	g	g	t	cc	at	tg	534
73	Ile	Val	Ser	Glu	Trp	Met	Gly	Tyr	Gly	Leu	Leu	His	Glu	Ser	Met	Leu						
74																	155	160	165			
76	agc	tcc	gt	ctc	cac	gc	cg	ca	cc	a	ag	t	gg	ct	g	ag	gg	gt	ct		582	
77	Ser	Ser	Val	Leu	His	Ala	Arg	Thr	Lys	Trp	Leu	Lys	Glu	Gly	Gly	Leu						
78																	170	175	180			
80	ctc	ctg	ccg	gc	tcc	gc	g	ag	ctc	t	tc	at	gc	cc	at	ag	gc	ca		630		
81	Leu	Leu	Pro	Ala	Ser	Ala	Glu	Leu	Phe	Ile	Ala	Pro	Ile	Ser	Asp	Gln						
82																	185	190	195			
84	atg	ctg	gaa	tgg	cg	ctg	gg	t	tc	t	gg	ag	ct	g	aa	g	ca	ca	tc	at	678	
85	Met	Leu	Glu	Trp	Arg	Leu	Gly	Phe	Trp	Ser	Gln	Val	Lys	Gln	His	Tyr						
86																	200	205	210	215		
88	gg	gt	g	ac	at	g	tc	ct	g	g	gc	t	tc	cc	at	g	cc	tg	tc	at	726	
89	Gly	Val	Asp	Met	Ser	Cys	Leu	Glu	Gly	Phe	Ala	Thr	Arg	Cys	Leu	Met						
90																	220	225	230			
92	gg	ca	tc	g	ag	at	c	tt	gt	ca	gg	tt	cc	gg	ca	gt	tg	ct		774		
93	Gly	His	Ser	Glu	Ile	Val	Val	Gln	Gly	Leu	Ser	Gly	Glu	Asp	Val	Leu						
94																	235	240	245			
96	gg	cc	cc	ca	cg	cg	tt	g	ct	ca	gg	tc	cc	gg	cc	gg	tt			822		
97	Ala	Arg	Pro	Gln	Arg	Phe	Ala	Gln	Leu	Glu	Leu	Ser	Arg	Ala	Gly	Leu						
98																	250	255	260			
100	gag	ca	g	g	ct	g	g	gg	gt	gg	gg	cg	cc	tt	cc	cg	tc	cg	tg	ac	870	
101	Glu	Gln	Glu	Leu	Glu	Ala	Gly	Val	Gly	Gly	Arg	Phe	Arg	Cys	Ser	Cys						
102																	265	270	275			
104	tat	gg	tc	g	cg	cc	at	g	tc	ttt	gg	at	tc	t	gg	tt	ca	g	tg	ac	918	
105	Tyr	Gly	Ser	Ala	Pro	Met	His	Gly	Phe	Ala	Ile	Trp	Phe	Gln	Val	Thr						
106																	280	285	290	295		
108	ttc	c	c	ttt	gg	gg	g	ag	tc	g	aa	cc	ct	gt	ct	tcc	acc	tc	cc		966	
109	Phe	Pro	Gly	Gly	Glu	Ser	Glu	Lys	Pro	Leu	Val	Leu	Ser	Thr	Ser	Pro						
110																	300	305	310			
112	ttt	cac	cc	g	cc	ac	cac	tgg	aaa	ca	gg	cg	ct	tg	tc	cc	ac	tc	cc	tt	1014	
113	Phe	His	Pro	Ala	Thr	His	Trp	Lys	Gln	Ala	Leu	Leu	Tyr	Leu	Asn	Glu						
114																	315	320	325			
116	cc	gt	ca	gt	g	ag	ca	g	ac	g	tt	tca	gg	ga	g	at	ac	cg	tg	ac	1062	
117	Pro	Val	Gln	Val	Glu	Gln	Asp	Thr	Asp	Val	Ser	Gly	Glu	Ile	Thr	Leu						
118																	330	335	340			
120	ct	cc	cc	cc	cc	cc	cc	cc	cg	cg	cg	gt	ct	tg	ct	tg	cc	tg	ac	ta	1110	
121	Leu	Pro	Ser	Arg	His	Asn	Pro	Arg	Arg	Leu	Arg	Val	Leu	Leu	Arg	Tyr						
122																	345	350	355			
124	aaa	gt	g	ga	g	ca	g	ag	g	ag	a	aa	g	ac	ttt	gg	at	g	ag	ac	1158	
125	Lys	Val	Gly	Asp	Gln	Glu	Lys	Thr	Lys	Asp	Phe	Ala	Met	Glu	Asp							
126																	360	365	370	375		
128	tga	gcgttgccaa	ggggcgaattc	gtttaaacct	gcaggactag																1201	

TECH CENTER 1600/2900

APR 08 2002

RECEIVED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/970,638

DATE: 03/27/2002  
TIME: 13:58:40

Input Set : F:\38155-20039.txt  
Output Set: N:\CRF3\03272002\I970638.raw

```

129 *
133 <210> SEQ ID NO: 2
134 <211> LENGTH: 375
135 <212> TYPE: PRT
136 <213> ORGANISM: Homo sapiens
138 <400> SEQUENCE: 2
139 Met Ser Gln Pro Lys Lys Arg Lys Leu Glu Ser Gly Gly Gly Glu
140 1 5 10 15
141 Gly Gly Glu Gly Thr Glu Glu Asp Gly Ala Glu Arg Glu Ala Ala
142 20 25 30
143 Leu Glu Arg Pro Arg Arg Thr Lys Arg Glu Arg Asp Gln Leu Tyr Tyr
144 35 40 45
145 Glu Cys Tyr Ser Asp Val Ser Val His Glu Glu Met Ile Ala Asp Arg
146 50 55 60
147 Val Arg Thr Asp Ala Tyr Arg Leu Gly Ile Leu Arg Asn Trp Ala Ala
148 65 70 75 80
149 Leu Arg Gly Lys Thr Val Leu Asp Val Gly Ala Gly Thr Gly Ile Leu
150 85 90 95
151 Ser Ile Phe Cys Ala Gln Ala Gly Ala Arg Arg Val Tyr Ala Val Glu
152 100 105 110
153 Ala Ser Ala Ile Trp Gln Gln Ala Arg Glu Val Val Arg Phe Asn Gly
154 115 120 125
155 Leu Glu Asp Arg Val His Val Leu Pro Gly Pro Val Glu Thr Val Glu
156 130 135 140
157 Leu Pro Glu Gln Val Asp Ala Ile Val Ser Glu Trp Met Gly Tyr Gly
158 145 150 155 160
159 Leu Leu His Glu Ser Met Leu Ser Ser Val Leu His Ala Arg Thr Lys
160 165 170 175
161 Trp Leu Lys Glu Gly Leu Leu Leu Pro Ala Ser Ala Glu Leu Phe
162 180 185 190
163 Ile Ala Pro Ile Ser Asp Gln Met Leu Glu Trp Arg Leu Gly Phe Trp
164 195 200 205
165 Ser Gln Val Lys Gln His Tyr Gly Val Asp Met Ser Cys Leu Glu Gly
166 210 215 220
167 Phe Ala Thr Arg Cys Leu Met Gly His Ser Glu Ile Val Val Gln Gly
168 225 230 235 240
169 Leu Ser Gly Glu Asp Val Leu Ala Arg Pro Gln Arg Phe Ala Gln Leu
170 245 250 255
171 Glu Leu Ser Arg Ala Gly Leu Glu Gln Glu Leu Glu Ala Gly Val Gly
172 260 265 270
173 Gly Arg Phe Arg Cys Ser Cys Tyr Gly Ser Ala Pro Met His Gly Phe
174 275 280 285
175 Ala Ile Trp Phe Gln Val Thr Phe Pro Gly Gly Glu Ser Glu Lys Pro
176 290 295 300
177 Leu Val Leu Ser Thr Ser Pro Phe His Pro Ala Thr His Trp Lys Gln
178 305 310 315 320
179 Ala Leu Leu Tyr Leu Asn Glu Pro Val Gln Val Glu Gln Asp Thr Asp
180 325 330 335
181 Val Ser Gly Glu Ile Thr Leu Leu Pro Ser Arg His Asn Pro Arg Arg

```

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/970,638

DATE: 03/27/2002  
TIME: 13:58:40

Input Set : F:\38155-20039.txt  
Output Set: N:\CRF3\03272002\I970638.raw

182	340	345	350		
183	Leu Arg Val	Leu Leu Arg Tyr	Lys Val Gly Asp Gln	Glu Lys Thr	
184	355	360	365		
185	Lys Asp Phe Ala Met	Glu Asp			
186	370	375			
188	<210> SEQ ID NO: 3				
189	<211> LENGTH: 1128				
190	<212> TYPE: DNA				
191	<213> ORGANISM: Homo sapiens				
193	<400> SEQUENCE: 3				
194	atgtcgca	ccaaagaaaag aaagctttag	tcggggggcg gcgccgaagg	aggggaggga 60	
195	actgaagagg	aagatggcgc ggagcggag	gccccctgg agcgaccccg	gaggactaag 120	
196	cgggaacggg	accagctgt	tacgcgtgc tactcgacg	tttcggtcca cgaggagatg 180	
197	atcgccggacc	gcgtccgcac	cgatgcctac cgcctggta	tccttcggaa ctgggcagca 240	
198	ctgcgaggca	agacggta	ggacgtggc gcccggcaccg	gcattctgag catcttctgt 300	
199	gcccaggccg	ggcccccggc	cgtgtacgcg	gtagaggcca gcccacatcg gcaacaggcc 360	
200	cgggaggtgg	tgcggttcaa	cgggctggag	gaccgggtgc acgtcctgcc gggaccagtg 420	
201	gagactgtag	agttgccgga	acaggtggat	gccatcgta gcgagtggat gggctacgga 480	
202	ctccctgcacg	agtccatgct	gagctccgtc	ctccacgcgc gaaccaagtg gctgaaggag 540	
203	gcccgtcttc	tcctgcggc	ctccgcgcag	ctcttcatag ccccatcg cgaccagatg 600	
204	ctggaatg	gcctgggctt	ctggagccag	gtgaagcagc actatggtgt ggacatgagc 660	
205	tgcctggagg	gcttcgcccac	gctgtgtctc	atgggcaact cggagatcg tgcagtgatg 720	
206	ttgtccggcg	aggacgtgct	ggcccgccg	cagcgcttg ctca	gctctccgc 780
207	gcccgttgg	agcaggagct	ggaggccgga	gtggccggc gcttcgctg cagctgttat 840	
208	ggctcggcgc	ccatgcattgg	cttgcattc	tggccagg tgaccttccc tggaggggag 900	
209	tcggagaaac	ccctggtgct	gtccacctcg	cctttcacc cggccactca ctggaaacag 960	
210	gcgctcctct	accta	acgca	gcccgtgcgtt gacggacgt ttca	gaggag 1020
211	atcacgctgc	tgcctcccg	gcacaacccc	cgtcgcctgc gcgtgctgct ggcataaaaa 1080	
212	gtgggagacc	aggaggagaa	gaccaaagac	tttgcattgg aggactga 1128	
214	<210> SEQ ID NO: 4				
215	<211> LENGTH: 304				
216	<212> TYPE: PRT				
217	<213> ORGANISM: Artificial Sequence				
219	<220> FEATURE:				
220	<223> OTHER INFORMATION: Consensus amino acid sequence				
222	<400> SEQUENCE: 4				
223	Ile Pro Thr Ser Tyr Asp Ile Ile Gly Asp Ile Val Ile Leu Asn Ile				
224	1	5	10	15	
225	Pro Asp Glu Leu Arg Pro Tyr Arg Lys Val Ile Gly Glu Ala Ile Leu				
226	20	25	30		
227	Lys Lys Thr Lys Lys Asn Glu Met Ala Val Val Lys Thr Val Leu Arg				
228	35	40	45		
229	Lys Gly Gly Glu Glu Asp Phe Glu Gly Leu Leu Gly Glu Gly Tyr Gly				
230	50	55	60		
231	Ile Thr Gly Glu Tyr Arg Thr Arg Arg Glu Leu Leu Ala Gly Glu Lys				
232	65	70	75	80	
233	Asp Ser Gly Glu Thr Val Thr Ile His Lys Glu Asn Gly Cys Arg Phe				
234	85	90	95		
235	Lys Leu Asp Val Ser Lys Val Tyr Phe Ser Pro Arg Leu Glu Gly Glu				

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/970,638

DATE: 03/27/2002  
TIME: 13:58:41

Input Set : F:\38155-20039.txt  
Output Set: N:\CRF3\03272002\I970638.raw

236	100	105	110	
237	Arg Glu Arg Leu Ala Lys Leu Val Lys Glu Gly Glu Val Val Val Asp			
238	115	120	125	
239	Met Phe Ala Gly Ile Gly Pro Phe Ser Ile Pro Ile Ala Lys His Ser			
240	130	135	140	
241	Lys Ala Lys Arg Val Tyr Ala Val Asp Leu Asn Pro Glu Ala Val Lys			
242	145	150	155	160
243	Tyr Leu Lys Glu Asn Ile Lys Leu Asn Lys Val Glu Asp Gln Pro Lys			
244	165	170	175	
245	Ile Gln Thr Val Val Asn Lys Ile Ile Pro Ile Leu Gly Asp Val Arg			
246	180	185	190	
247	Glu Val Ile Pro Glu Lys Glu Ala Gly Val His Ser Leu Val Lys Leu			
248	195	200	205	
249	Val Val Ala Asp Arg Val Ile Met Pro Leu Pro Leu Ala Arg Val Gly			
250	210	215	220	
251	Gln Ile Ser Ala His Glu Phe Leu Asp Lys Ala Leu Arg Lys Val Lys			
252	225	230	235	240
253	Asp Gly Gly Ile Leu Val Ser Val Ile His Tyr Tyr Glu Phe Leu Pro			
254	245	250	255	
255	Glu Lys Ala Ile Pro Phe Leu Pro Val Glu Arg Ile Lys Arg Ala Ala			
256	260	265	270	
257	Glu Lys Glu Gly Val Lys Val Glu Ile Leu Lys Lys Arg Lys Val Lys			
258	275	280	285	
259	Asn Tyr Ala Pro Gly Val Tyr His Val Val Leu Asp Ala Arg Val Asn			
260	290	295	300	
262	<210> SEQ ID NO: 5			
263	<211> LENGTH: 185			
264	<212> TYPE: PRT			
265	<213> ORGANISM: Artificial Sequence			
267	<220> FEATURE:			
268	<223> OTHER INFORMATION: Consensus amino acid sequence			
270	<400> SEQUENCE: 5			
271	Pro Asp Lys Cys Thr Leu His Leu Thr Ala Ile Glu Asp Thr Gln Tyr			
272	1	5	10	15
273	Lys Asp Glu Lys Ile Asn Phe Trp Asp Asp Val Tyr Gly Phe Asx Met			
274	20	25	30	
275	Ser Cys Met Lys Asp Ala Ala Ile Lys Glu Pro Ile Val Asp Pro Lys			
276	35	40	45	
277	Tyr Asn His Val Val Asp Pro Asn Gln Ile Ile Thr Asp Pro Cys Thr			
278	50	55	60	
279	Ile Lys Glu Phe Asp Met His Thr Val Lys Ile Ser Asp Leu Glu Thr			
280	65	70	75	80
281	Phe Thr Ser Pro Phe Lys Leu Thr Val Lys Arg Asn Gly Met Leu His			
282	85	90	95	
283	Ala Phe Thr Ala Trp Phe Asp Val Tyr Phe Thr Ser Lys Arg Cys His			
284	100	105	110	
285	Pro Ile Pro Glu Ala Ile Asp Lys Pro Val Gly Phe Ser Thr Gly Pro			
286	115	120	125	
287	Gln Ser Pro Tyr Thr His Trp Lys Gln Thr Ile Phe Tyr Met Glu Asp			

VERIFICATION SUMMARY  
PATENT APPLICATION: US/09/970,638

DATE: 03/27/2002  
TIME: 13:58:42

Input Set : F:\38155-20039.txt  
Output Set: N:\CRF3\03272002\I970638.raw

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date